

1641



ENTERED

IFWO

RAW SEQUENCE LISTING
PATENT APPLICATION: US/10/644,875

DATE: 01/21/2004
TIME: 14:50:31

Input Set : N:\CrF3\RULE60\10644875.RAW.txt
Output Set: N:\CRF4\01212004\J644875.raw

SEQUENCE LISTING

- 3 (1) GENERAL INFORMATION:
 5 (i) APPLICANT: Wei, et al.
 7 (ii) TITLE OF INVENTION: Transforming Growth Factor Alpha HII
 9 (iii) NUMBER OF SEQUENCES: 14
 11 (iv) CORRESPONDENCE ADDRESS:
 12 (A) ADDRESSEE: Human Genome Sciences, Inc.
 13 (B) STREET: 9410 Key West Ave
 14 (C) CITY: Rockville
 15 (D) STATE: MD
 16 (E) COUNTRY: US
 17 (F) ZIP: 20850
 19 (v) COMPUTER READABLE FORM:
 20 (A) MEDIUM TYPE: Floppy disk
 21 (B) COMPUTER: IBM PC compatible
 22 (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 23 (D) SOFTWARE: PatentIn Release #1.0, Version #1.30
 25 (vi) CURRENT APPLICATION DATA:
 C--> 26 (A) APPLICATION NUMBER: US/10/644,875
 C--> 27 (B) FILING DATE: 21-Aug-2003
 28 (C) CLASSIFICATION: 536
 C--> 37 (vii) PRIOR APPLICATION DATA:
 31 (A) APPLICATION NUMBER: US/09/227,853A
 32 (B) FILING DATE: 11-JAN-1999
 34 (A) APPLICATION NUMBER: 08/930,564
 35 (B) FILING DATE: 30-JAN-1998
 38 (A) APPLICATION NUMBER: WO PCT/US95/06386
 39 (B) FILING DATE: 19-MAY-1995
 C--> 41 (viii) ATTORNEY/AGENT INFORMATION:
 42 (A) NAME: Brookes, A. Anders
 43 (B) REGISTRATION NUMBER: 36,373
 44 (C) REFERENCE/DOCKET NUMBER: PF174USD1
 C--> 46 (ix) TELECOMMUNICATION INFORMATION:
 47 (A) TELEPHONE: 301-309-8504
 48 (B) TELEFAX: 301-309-8439
 51 (2) INFORMATION FOR SEQ ID NO: 1:
 53 (i) SEQUENCE CHARACTERISTICS:
 54 (A) LENGTH: 1695 base pairs
 55 (B) TYPE: nucleic acid
 56 (C) STRANDEDNESS: single
 57 (D) TOPOLOGY: linear
 59 (ii) MOLECULE TYPE: DNA (genomic)
 62 (ix) FEATURE:

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/644,875

DATE: 01/21/2004

TIME: 14:50:31

Input Set : N:\Crf3\RULE60\10644875.RAW.txt

Output Set: N:\CRF4\01212004\J644875.raw

63 (A) NAME/KEY: CDS
 64 (B) LOCATION: 323..1444
 68 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:
 70 CACTCGTCTG CCCCTGGACT CCCGTCTCCT CCTGTCCTCC GGCTTCCCAG AGCTCCCTCC 60
 72 TTATGGCAGC AGCTTCCCGC GTCTCCGGCG CAGTTCTCAG CGGACGACCC TCTCGCTCCG 120
 74 GGGCTGAGCC CAGTCCCTGG ATGTTGCTGA AACTCTCGAG ATCATGCCCG GGTTTGGCTG 180
 76 CTGCTTCCCC GCCGGGTGCC ACTGCCACCG CGGCCGCCTC TGCTGCCGCC GTCCGCCGGGA 240
 78 TGCTCAGTAG CCCGCTGCC CGGCCGCCGG ATCCTGTGTT CCTCGGAAGC CGTTTGCTGC 300
 80 TGCAGAGTTG CACGAAGTAG TC ATG GTG CTG TGG GAG TCC CCG CGG CAG TGC 352
 81 Met Val Leu Trp Glu Ser Pro Arg Gln Cys
 82 1 5 10
 84 AGC AGC TGG ACA CTT TGC GAG GGC TTT TGC TGG CTG CTG CTG CTG CCC 400
 85 Ser Ser Trp Thr Leu Cys Glu Gly Phe Cys Trp Leu Leu Leu Pro
 86 15 20 25
 88 GTC ATG CTA CTC ATC GTA GCC CGC CCG GTG AAG CTC GCT GCT TTC CCT 448
 89 Val Met Leu Leu Ile Val Ala Arg Pro Val Lys Leu Ala Ala Phe Pro
 90 30 35 40
 92 ACC TCC TTA AGT GAC TGC CAA ACG CCC ACC GGC TGG AAT TGC TCT GGT 496
 93 Thr Ser Leu Ser Asp Cys Gln Thr Pro Thr Gly Trp Asn Cys Ser Gly
 94 45 50 55
 96 TAT GAT GAC AGA GAA AAT GAT CTC TTC CTC TGT GAC ACC AAC ACC TGT 544
 97 Tyr Asp Asp Arg Glu Asn Asp Leu Phe Leu Cys Asp Thr Asn Thr Cys
 98 60 65 70
 100 AAA TTT GAT GGG GAA TGT TTA AGA ATT GGA GAC ACT GTG ACT TGC GTC 592
 101 Lys Phe Asp Gly Glu Cys Leu Arg Ile Gly Asp Thr Val Thr Cys Val
 102 75 80 85 90
 104 TGT CAG TTC AAG TGC AAC AAT GAC TAT GTG CCT GTG TGT GGC TCC AAT 640
 105 Cys Gln Phe Lys Cys Asn Asn Asp Tyr Val Pro Val Cys Gly Ser Asn
 106 95 100 105
 108 GGG GAG AGC TAC CAG AAT GAG TGT TAC CTG CGA CAG GCT GCA TGC AAA 688
 109 Gly Glu Ser Tyr Gln Asn Glu Cys Tyr Leu Arg Gln Ala Ala Cys Lys
 110 110 115 120
 112 CAG CAG AGT GAG ATA CTT GTG GTG TCA GAA GGA TCA TGT GCC ACA GAT 736
 113 Gln Gln Ser Glu Ile Leu Val Val Ser Glu Gly Ser Cys Ala Thr Asp
 114 125 130 135
 116 GCA GGA TCA GGA TCT GGA GAT GGA GTC CAT GAA GGC TCT GGA GAA ACT 784
 117 Ala Gly Ser Gly Ser Gly Asp Gly Val His Glu Gly Ser Gly Glu Thr
 118 140 145 150
 120 AGT CAA AAG GAG ACA TCC ACC TGT GAT ATT TGC CAG TTT GGT GCA GAA 832
 121 Ser Gln Lys Glu Thr Ser Thr Cys Asp Ile Cys Gln Phe Gly Ala Glu
 122 155 160 165 170
 124 TGT GAC GAA GAT GCC GAG GAT GTC TGG TGT GTG TGT AAT ATT GAC TGT 880
 125 Cys Asp Glu Asp Ala Glu Asp Val Trp Cys Val Cys Asn Ile Asp Cys
 126 175 180 185
 128 TCT CAA ACC AAC TTC AAT CCC CTC TGC GCT TCT GAT GGG AAA TCT TAT 928
 129 Ser Gln Thr Asn Phe Asn Pro Leu Cys Ala Ser Asp Gly Lys Ser Tyr
 130 190 195 200
 132 GAT AAT GCA TGC CAA ATC AAA GAA GCA TCG TGT CAG AAA CAG GAG AAA 976
 133 Asp Asn Ala Cys Gln Ile Lys Glu Ala Ser Cys Gln Lys Gln Glu Lys

RAW SEQUENCE LISTING
PATENT APPLICATION: US/10/644,875

DATE: 01/21/2004
TIME: 14:50:31

Input Set : N:\Crf3\RULE60\10644875.RAW.txt
Output Set: N:\CRF4\01212004\J644875.raw

134	205	210	215	
136	ATT GAA GTC ATG TCT TTG GGT CGA TGT CAA GAT AAC ACA ACT ACA ACT			1024
137	Ile Glu Val Met Ser Leu Gly Arg Cys Gln Asp Asn Thr Thr Thr Thr			
138	220	225	230	
140	ACT AAG TCT GAA GAT GGG CAT TAT GCA AGA ACA GAT TAT GCA GAG AAT			1072
141	Thr Lys Ser Glu Asp Gly His Tyr Ala Arg Thr Asp Tyr Ala Glu Asn			
142	235	240	245	250
144	GCT AAC AAA TTA GAA GAA AGT GCC AGA GAA CAC CAC ATA CCT TGT CCG			1120
145	Ala Asn Lys Leu Glu Glu Ser Ala Arg Glu His His Ile Pro Cys Pro			
146	255	260	265	
148	GAA CAT TAC AAT GGC TTC TGC ATG CAT GGG AAG TGT GAG CAT TCT ATC			1168
149	Glu His Tyr Asn Gly Phe Cys Met His Gly Lys Cys Glu His Ser Ile			
150	270	275	280	
152	AAT ATG CAG GAG CCA TCT TGC AGG TGT GAT GCT GGT TAT ACT GGA CAA			1216
153	Asn Met Gln Glu Pro Ser Cys Arg Cys Asp Ala Gly Tyr Thr Gly Gln			
154	285	290	295	
156	CAC TGT GAA AAA AAG GAC TAC AGT GTT CTA TAC GTT GTT CCC GGT CCT			1264
157	His Cys Glu Lys Lys Asp Tyr Ser Val Leu Tyr Val Val Pro Gly Pro			
158	300	305	310	
160	GTA CGA TTT CAG TAT GTC TTA ATC GCA GCT GTG ATT GGA ACA ATT CAG			1312
161	Val Arg Phe Gln Tyr Val Leu Ile Ala Ala Val Ile Gly Thr Ile Gln			
162	315	320	325	330
164	ATT GCT GTC ATC TGT GTG GTG GTC CTC TGC ATC ACA AGG AAA TGC CCC			1360
165	Ile Ala Val Ile Cys Val Val Leu Cys Ile Thr Arg Lys Cys Pro			
166	335	340	345	
168	AGA AGC AAC AGA ATT CAC AGA CAG AAG CAA AAT ACA GGG CAC TAC AGT			1408
169	Arg Ser Asn Arg Ile His Arg Gln Lys Gln Asn Thr Gly His Tyr Ser			
170	350	355	360	
172	TCG GAC AAT ACA ACA AGA GCG TCC ACG AGG TTA ATC TAAAGGGAGC			1454
173	Ser Asp Asn Thr Thr Arg Ala Ser Thr Arg Leu Ile			
174	365	370		
176	ATGTTTCACA GTGGCTGGAC TACCGAGAGC TTGGACTACA CAATACAGTA TTATAGACAA			1514
178	AAGAATAAGA CAAGAGATCT ACACATGTTG CCTTCATTT GTGGTAATCT ACACCAATGA			1574
180	AAACATGTAC TACAGCTATA TTTGATTATG TATGGATATA TTTGAAATAG TATACATTGT			1634
182	CTTGATGTTT TTTCTGTAAT GTAAATAAAC TATTTATATC ACACAAAAAA AAAAAAAA			1694
184	A			1695
187	(2) INFORMATION FOR SEQ ID NO: 2:			
189	(i) SEQUENCE CHARACTERISTICS:			
190	(A) LENGTH: 374 amino acids			
191	(B) TYPE: amino acid			
192	(D) TOPOLOGY: linear			
194	(ii) MOLECULE TYPE: protein			
196	(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:			
198	Met Val Leu Trp Glu Ser Pro Arg Gln Cys Ser Ser Trp Thr Leu Cys			
199	1	5	10	15
201	Glu Gly Phe Cys Trp Leu Leu Leu Pro Val Met Leu Leu Ile Val			
202	20	25	30	
204	Ala Arg Pro Val Lys Leu Ala Ala Phe Pro Thr Ser Leu Ser Asp Cys			
205	35	40	45	

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/644,875

DATE: 01/21/2004

TIME: 14:50:31

Input Set : N:\Crf3\RULE60\10644875.RAW.txt

Output Set: N:\CRF4\01212004\J644875.raw

207 Gln Thr Pro Thr Gly Trp Asn Cys Ser Gly Tyr Asp Asp Arg Glu Asn
 208 50 55 60
 210 Asp Leu Phe Leu Cys Asp Thr Asn Thr Cys Lys Phe Asp Gly Glu Cys
 211 65 70 75 80
 213 Leu Arg Ile Gly Asp Thr Val Thr Cys Val Cys Gln Phe Lys Cys Asn
 214 85 90 95
 216 Asn Asp Tyr Val Pro Val Cys Gly Ser Asn Gly Glu Ser Tyr Gln Asn
 217 100 105 110
 219 Glu Cys Tyr Leu Arg Gln Ala Ala Cys Lys Gln Gln Ser Glu Ile Leu
 220 115 120 125
 222 Val Val Ser Glu Gly Ser Cys Ala Thr Asp Ala Gly Ser Gly Ser Gly
 223 130 135 140
 225 Asp Gly Val His Glu Gly Ser Gly Glu Thr Ser Gln Lys Glu Thr Ser
 226 145 150 155 160
 228 Thr Cys Asp Ile Cys Gln Phe Gly Ala Glu Cys Asp Glu Asp Ala Glu
 229 165 170 175
 231 Asp Val Trp Cys Val Cys Asn Ile Asp Cys Ser Gln Thr Asn Phe Asn
 232 180 185 190
 234 Pro Leu Cys Ala Ser Asp Gly Lys Ser Tyr Asp Asn Ala Cys Gln Ile
 235 195 200 205
 237 Lys Glu Ala Ser Cys Gln Lys Gln Glu Lys Ile Glu Val Met Ser Leu
 238 210 215 220
 240 Gly Arg Cys Gln Asp Asn Thr Thr Thr Thr Lys Ser Glu Asp Gly
 241 225 230 235 240
 243 His Tyr Ala Arg Thr Asp Tyr Ala Glu Asn Ala Asn Lys Leu Glu Glu
 244 245 250 255
 246 Ser Ala Arg Glu His His Ile Pro Cys Pro Glu His Tyr Asn Gly Phe
 247 260 265 270
 249 Cys Met His Gly Lys Cys Glu His Ser Ile Asn Met Gln Glu Pro Ser
 250 275 280 285
 252 Cys Arg Cys Asp Ala Gly Tyr Thr Gly Gln His Cys Glu Lys Lys Asp
 253 290 295 300
 255 Tyr Ser Val Leu Tyr Val Val Pro Gly Pro Val Arg Phe Gln Tyr Val
 256 305 310 315 320
 258 Leu Ile Ala Ala Val Ile Gly Thr Ile Gln Ile Ala Val Ile Cys Val
 259 325 330 335
 261 Val Val Leu Cys Ile Thr Arg Lys Cys Pro Arg Ser Asn Arg Ile His
 262 340 345 350
 264 Arg Gln Lys Gln Asn Thr Gly His Tyr Ser Ser Asp Asn Thr Thr Arg
 265 355 360 365
 267 Ala Ser Thr Arg Leu Ile
 268 370
 270 (2) INFORMATION FOR SEQ ID NO: 3:
 272 (i) SEQUENCE CHARACTERISTICS:
 273 (A) LENGTH: 30 base pairs
 274 (B) TYPE: nucleic acid
 275 (C) STRANDEDNESS: single
 276 (D) TOPOLOGY: linear
 278 (ii) MOLECULE TYPE: DNA (genomic)

RAW SEQUENCE LISTING
PATENT APPLICATION: US/10/644,875

DATE: 01/21/2004
TIME: 14:50:31

Input Set : N:\Crf3\RULE60\10644875.RAW.txt
Output Set: N:\CRF4\01212004\J644875.raw

283 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:
285 CCCGGATCCG CACGAGACAT ACCTTGTCCG 30
287 (2) INFORMATION FOR SEQ ID NO: 4:
289 (i) SEQUENCE CHARACTERISTICS:
290 (A) LENGTH: 32 base pairs
291 (B) TYPE: nucleic acid
292 (C) STRANDEDNESS: single
293 (D) TOPOLOGY: linear
295 (ii) MOLECULE TYPE: DNA (genomic)
300 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:
302 GGGAAAGCTTT TAATACTGAA ATCGTACAGG AC 32
304 (2) INFORMATION FOR SEQ ID NO: 5:
306 (i) SEQUENCE CHARACTERISTICS:
307 (A) LENGTH: 33 base pairs
308 (B) TYPE: nucleic acid
309 (C) STRANDEDNESS: single
310 (D) TOPOLOGY: linear
312 (ii) MOLECULE TYPE: DNA (genomic)
317 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:
319 CGCGGGATCCG CCATCATGGT GCTGTGGGAG TCC 33
321 (2) INFORMATION FOR SEQ ID NO: 6:
323 (i) SEQUENCE CHARACTERISTICS:
324 (A) LENGTH: 31 base pairs
325 (B) TYPE: nucleic acid
326 (C) STRANDEDNESS: single
327 (D) TOPOLOGY: linear
329 (ii) MOLECULE TYPE: DNA (genomic)
334 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:
336 GCGTCTAGAC TAGTATAGAA CACTGTAGTC C 31
338 (2) INFORMATION FOR SEQ ID NO: 7:
340 (i) SEQUENCE CHARACTERISTICS:
341 (A) LENGTH: 31 base pairs
342 (B) TYPE: nucleic acid
343 (C) STRANDEDNESS: single
344 (D) TOPOLOGY: linear
346 (ii) MOLECULE TYPE: DNA (genomic)
351 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:
353 GCGTCTAGAC TAGTATAGAA CACTGTAGTC C 31
355 (2) INFORMATION FOR SEQ ID NO: 8:
357 (i) SEQUENCE CHARACTERISTICS:
358 (A) LENGTH: 31 base pairs
359 (B) TYPE: nucleic acid
360 (C) STRANDEDNESS: single
361 (D) TOPOLOGY: linear
363 (ii) MOLECULE TYPE: DNA (genomic)
368 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:
370 GCGTCTAGAC TAGTATAGAA CACTGTAGTC C 31
372 (2) INFORMATION FOR SEQ ID NO: 9:
374 (i) SEQUENCE CHARACTERISTICS:

VERIFICATION SUMMARY

PATENT APPLICATION: US/10/644,875

DATE: 01/21/2004

TIME: 14:50:32

Input Set : N:\Crf3\RULE60\10644875.RAW.txt

Output Set: N:\CRF4\01212004\J644875.raw

L:26 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]
L:27 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]
L:37 M:220 C: Keyword misspelled or invalid format, [(vii) PRIOR APPLICATION DATA:]
L:41 M:220 C: Keyword misspelled or invalid format, [(viii) ATTORNEY/AGENT INFORMATION:]
L:46 M:220 C: Keyword misspelled or invalid format, [(ix) TELECOMMUNICATION INFORMATION:]